Scalar and Vector Partitions of the Probability Score: Part I. Two-State Situation[1]

ALLAN H. MURPHY[2]

Dept. of Meteorology and Oceanography, University of Michigan, Ann Arbor 80302

ABSTRACT

Scalar and vector partitions of the probability score (PS) in the two-state situation are described and compared. These partitions, which are based upon expressions for the PS in which probability forecasts are considered to be scalars and vectors, respectively, provide similar, but not equivalent (i.e., linearly related), measures of the reliability and resolution of the forecasts. Specifically, the reliability (resolution) of the forecasts according to the scalar partition is, in general, greater (less) than their reliability (resolution) according to the vector partition. A sample collection of forecasts is used to illustrate the differences between these partitions.

Several questions related to the use of scalar and vector partitions of the PS in the two-state situation are discussed, including the interpretation of the results of previous forecast evaluation studies and the relative merits of these partitions. The discussions indicate that the partition most often used in such studies has been a special "scalar" partition, a partition which is equivalent to the vector partition in the two-state situation, and that the vector partition is more appropriate than the scalar partition.

1. Introduction

Sanders (1958, 1963) demonstrated that the probability score (PS) (Brier, 1950) for a collection of probability forecasts, in which the probabilities can assume only a finite set of values, can be partitioned into two terms, each of which is a measure of a particular attribute of the forecasts. Sanders' partition is based upon an expression for the PS in which a probability forecast is considered to be a *scalar* quantity, i.e., in which each probability is considered to be a separate forecast (Sanders, 1963, p. 192).[3]

In reality, a probability forecast consists of a set of two or more probabilities. Thus, we believe that such a partition should be based upon an expression for the PS in which a probability forecast is considered to be a *vector* quantity. The purposes of this paper are to describe and compare scalar and vector partitions of the

---

[3] Sanders states that "we have chosen to consider each individual probability statement as a separate forecast." However, in two-state, i.e., "precipitation–no precipitation" situations, Sanders actually considers only the probabilities assigned to one of the two states. These considerations lead to the formulation of two different partitions, the scalar partition and the special "scalar" partition, respectively (see Sections 3, 6 and 7).

PS and to briefly discuss several questions related to the use of these partitions. In this paper we consider only the two-state ($N = 2$) situation. We shall describe and compare scalar and vector partitions of the PS in N-state ($N > 2$) situations in a separate paper (see Murphy, 1971b).

In Section 2 we describe the differences between scalar and vector forecasts and observations and introduce notation to identify these quantities. We obtain expressions for the scalar and vector partitions of the PS in Sections 3 and 4, respectively. In Section 5 we compare the scalar and vector partitions and demonstrate that these partitions are not equivalent, i.e., linearly related. A sample collection of forecasts is used to illustrate the differences between these partitions in Section 6. Several questions related to the use of such partitions are discussed in Section 7, including the interpretation of the results of previous forecast evaluation studies and the relative merits of these partitions. Section 8 consists of a brief summary and conclusion.

2. Scalar and vector forecasts and observations

We assume that the forecasts and observations relate to situations in which the range of the variable of concern has been divided into a set of two mutually exclusive and collectively exhaustive states $\{s_1, s_2\}$. Specifically, we assume that the variable of concern is
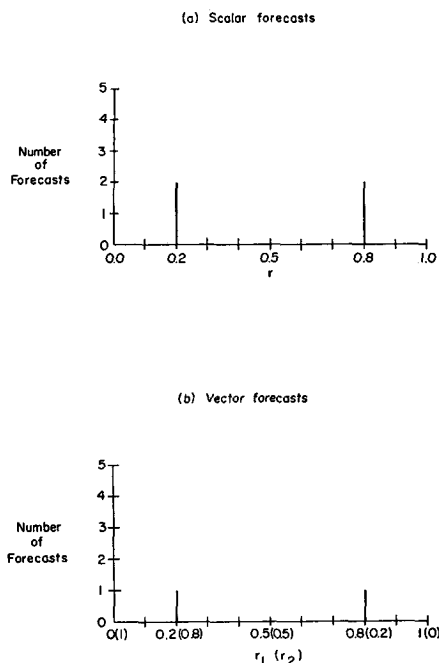
(a) Scalar forecasts

(b) Vector forecasts

FIG. 1. Two occasions on which the forecast probabilities of precipitation are 0.2 and 0.8, respectively, depicted within the appropriate framework when forecasts are considered to be (a) scalars and (b) vectors.

precipitation amount and that the two states are "precipitation" (s_1) and "no precipitation" (s_2).

When the probability assigned to each state on each occasion is assumed to constitute a forecast, we denote the forecast by a scalar r $(0 \leqslant r \leqslant 1)$ and the relevant observation by a scalar d, where d equals one if the state of concern occurs and zero otherwise. On the other hand, when the probabilities assigned to the set of states on each occasion are assumed to constitute a forecast, we denote the forecast by a row vector

$$\mathbf{r} = (r_1, r_2)(r_n \geqslant 0, \sum_n r_n = 1; n = 1,2),$$

and the relevant observation by a row vector $\mathbf{d} = (d_1, d_2)$, where d_n equals one if state s_n occurs and zero otherwise $(n = 1, 2)$.

In order to indicate the differences between scalar and vector forecasts, we consider two occasions on which the forecast probabilities of precipitation are 0.2 and 0.8, respectively. In the scalar framework we would have *four* forecasts (r), two for which the probability (r) is 0.2, one relating to precipitation and one relating to no precipitation, and two for which the probability (r) is 0.8, one relating to precipitation and one relating to no precipitation. On the other hand, in the vector framework we would have only *two* forecasts $(\mathbf{r})$, one for which the probabilities of precipitation (r_1) and no precipitation (r_2) are 0.2 and 0.8, respectively, and one for which the probabilities of precipitation (r_1)

and no precipitation (r_2) are 0.8 and 0.2, respectively.

We can also indicate the differences between scalar and vector forecasts by describing the appropriate framework within which to depict the forecasts. When forecasts are considered to be scalars, they can be depicted simply as points on a unit line segment. Each point on this line segment corresponds to a particular scalar forecast (and vice versa), and the end points of the line segment represent the observations as well as the categorical forecasts.[4] When forecasts are considered to be vectors, they can be depicted within the framework of a regular $(N-1)$-dimensional simplex.[5] In the two-state $(N=2)$ situation the regular simplex can be represented as a unit line segment, in which a "length" coordinate system is used to depict the forecasts.[6,7] Each point on this line segment corresponds to a particular vector forecast (and vice versa), and the end points of the line segment represent the observations as well as the categorical forecasts.[8] The probabilities on the two occasions of concern are depicted as scalar and vector forecasts in Figs. 1a and 1b, respectively.

The partitions of the PS of concern in this paper are based upon the assumption that the probabilities which constitute the forecasts can assume only a finite set of values. Specifically, we assume that the collection of forecasts of concern consists of M scalar or $K (=M/2)$ vector forecasts and that the probabilities can assume only S distinct values. Then, we can identify S distinct scalar forecasts and T distinct vector forecasts, where

[4] In the scalar framework the right-hand end point represents the occurrence, and the categorical forecast of the occurrence, of the state of concern (i.e., precipitation or no precipitation), while the left-hand end point represents the non-occurrence, and the categorical forecast of the non-occurrence, of the state of concern.

[5] A regular simplex is a line segment in the two-state $(N=2)$ situation, an equilateral triangle in the three-state $(N=3)$ situation, and a regular tetrahedron in the four-state $(N=4)$ situation (see Pontryagin, 1952, 10–12; see also Murphy, 1971a,b).

[6] In a regular simplex in which a "content" coordinate system is used to depict the forecasts, the perpendicular distances between the point (in the simplex) which represents a vector forecast and the "faces" of the simplex are equal to the components of the forecast. Thus, in the two-state $(N=2)$ situation, in which a content coordinate system becomes a length coordinate system, the distances between the point which represents a vector forecast and the end points of the unit line segment are the two components of the forecast (see Springer, 1964, 119–122).

[7] In the two-state $(N=2)$ situation, then, the unit line segment represents an appropriate framework within which to depict both scalar and vector forecasts. Scalar forecasts can always be depicted within this framework. On the other hand, the appropriate framework within which to depict vector forecasts is the $(N-1)$-dimensional simplex (see footnote 5). Thus, the differences between scalar and vector forecasts become more evident in N-state $(N>2)$ situations.

[8] In the vector framework the right-hand end point represents the occurrence, and the categorical forecast of the occurrence, of precipitation *and* the non-occurrence, and the categorical forecast of the non-occurrence, of no precipitation, while the left-hand end point represents the occurrence, and the categorical forecast of the occurrence, of no precipitation *and* the non-occurrence, and the categorical forecast of the non-occurrence, of precipitation.

T is equal to S in the two-state ($N=2$) situation.[9] Thus, we can identify S subcollections of scalar forecasts, where subcollection s consists of the M^s scalar forecasts for which

$$r_m = r^s \ (m=1, \ldots, M^s; \ \sum_s M^s = M; \ s=1, \ldots, S),$$

and T subcollections of vector forecasts, where subcollection t consists of the K^t vector forecasts for which

$$\mathbf{r}_k = \mathbf{r}^t \ (k=1, \ldots, K^t; \ \sum_t K^t = K; \ t=1, \ldots, T),$$

where $\mathbf{r}_k = (r_{1k}, r_{2k})$ and $\mathbf{r}^t = (r_1{}^t, r_2{}^t)$. For these subcollections we denote the relevant scalar observations by $d_m{}^s$ ($m=1, \ldots, M^s$) and the relevant vector observations by $\mathbf{d}_k{}^t$ ($k=1, \ldots, K^t$), where $\mathbf{d}_k{}^t = (d_{1k}{}^t, d_{2k}{}^t)$.

3. Scalar partition

The PS for a collection of M scalar forecasts r_m ($m=1, \ldots, M$) is $PS(r,d)$, where

$$PS(r,d) = (1/M) \sum_{m=1}^{M} (r_m - d_m)^2 \tag{1}$$

(Sanders, 1958, p. 38; Sanders, 1963, p. 192). Note that the range of $PS(r,d)$ is the closed unit interval $[0,1]$.

For the subcollection of M^s forecasts for which $r_m = r^s$, the PS is $PS^s(r,d)$, where

$$PS^s(r,d) = (1/M^s) \sum_{m=1}^{M^s} (r^s - d_m{}^s)^2,$$

or, since $d_m{}^s$ equals one or zero,

$$PS^s(r,d) = (r^s)^2 - 2r^s \Big[(1/M^s) \sum_{m=1}^{M^s} d_m{}^s \Big]$$
$$+ (1/M^s) \sum_{m=1}^{M^s} d_m{}^s. \tag{2}$$

[9] We use different superscripts to identify these distinct scalar and vector forecasts in the two-state ($N=2$) situation in order to ensure that the reader is able to distinguish between scalar forecasts and observations and the components of vector forecasts and observations. Moreover, while the number of distinct scalar forecasts is S in N-state ($N>2$) as well as in two-state situations, the number of distinct vector forecasts in N-state ($N \geqslant 2$) situations is T, where

$$T = \sum_{s=1}^{S} \binom{N+s-4}{s-1} (S-s+1),$$

in which $\binom{x}{y} = x! [y! (x-y)!]$ for $0 \geqslant y \geqslant x$, $\binom{x}{y} = 1$ for $x = -1$, $y=0$, and $\binom{x}{y} = 0$ otherwise. This expression is valid only if the set of S distinct scalar forecasts $r^s (s=1, \cdots, S)$ includes the values one and zero and if the difference between adjacent probability values is constant. For example, if this difference is 0.1 in a three-state ($N=3$) situation, then $S=11$ and $r^s = 0.0(0.1)1.0$ and, as a result, $T=66$ and $r^t = (1.0, 0.0, 0.0)$, $(0.9, 0.1, 0.0)$, $\cdots$, $(0.0, 0.0, 1.0)$. Note that, in general, T is greater than or equal to S and that T equals S only in the two-state situation.

Let

$$d_m{}^s = \bar{d}^s + (d_m{}^s)^0,$$

where

$$\bar{d}^s = (1/M^s) \sum_{m=1}^{M^s} d_m{}^s$$

and $(d_m{}^s)^0$ is the difference between $d_m{}^s$ and $\bar{d}^s$. If we substitute $\bar{d}^s$ into (2) and complete the square, then $PS^s(r,d)$ becomes

$$PS^s(r,d) = (r^s - \bar{d}^s)^2 + \bar{d}^s(1 - \bar{d}^s). \tag{3}$$

Therefore, $PS(r,d)$, the weighted sum of the subcollection scores, can be expressed as

$$PS(r,d) = (1/M) \sum_{s=1}^{S} M^s (r^s - \bar{d}^s)^2$$
$$+ (1/M) \sum_{s=1}^{S} M^s \bar{d}^s (1 - \bar{d}^s). \tag{4}$$

$PS(r,d)$, in (4), represents the scalar partition of the PS. Sanders (1958) referred to the two terms on the right-hand side (RHS) of (4) as measures of the attributes *reliability* and *resolution*, respectively.[10,11]

4. Vector partition

The PS for a collection of K ($=M/2$) vector forecasts $\mathbf{r}_k = (r_{1k}, r_{2k})$ ($k=1, \ldots, K$) is $PS(\mathbf{r}, \mathbf{d})$, where

$$PS(\mathbf{r}, \mathbf{d}) = (1/K) \sum_{k=1}^{K} \sum_{n=1}^{2} (r_{nk} - d_{nk})^2 \tag{5a}$$

(Brier, 1950, p. 1; Murphy and Epstein, 1967, p. 751), or, in vector notation,

$$PS(\mathbf{r}, \mathbf{d}) = (1/K) \sum_{k=1}^{K} (\mathbf{r}_k - \mathbf{d}_k)(\mathbf{r}_k - \mathbf{d}_k)', \tag{5b}$$

(see Winkler and Murphy, 1968, 751–752), where a prime denotes a column vector. Note that the range of $PS(\mathbf{r}, \mathbf{d})$ is the closed interval $[0,2]$.

For the subcollection of K^t forecasts for which $\mathbf{r}_k = \mathbf{r}^t$, the PS is $PS^t(\mathbf{r}, \mathbf{d})$, where

$$PS^t(\mathbf{r}, \mathbf{d}) = (1/K^t) \sum_{k=1}^{K^t} (\mathbf{r}^t - \mathbf{d}_k{}^t)(\mathbf{r}^t - \mathbf{d}_k{}^t)',$$

[10] Subsequently, Sanders (1963) referred to these attributes as *validity* and *sharpness*, respectively. We use the original terms because we prefer to define the term validity in a more general manner (see Murphy and Winkler, 1970, 281–282) and because the term sharpness, as defined by Bross (1953, 48–52), relates to the degree to which the forecast probabilities, rather than the observed relative frequencies, approach one or zero.

[11] Note that the measure of resolution depends directly upon the observations and only indirectly upon the forecasts. Thus, resolution is an attribute of the observations rather than of the forecasts. However, since this measure depends indirectly upon the forecasts, we shall continue to speak of the "resolution of the forecasts."

or

$$PS^t(\mathbf{r},\mathbf{d}) = \mathbf{r}^t(\mathbf{r}^t)' - 2\mathbf{r}^t\big[(1/K^t) \sum_{k=1}^{K^t} (\mathbf{d}_k{}^t)'\big]$$

$$+ (1/K^t) \sum_{k=1}^{K^t} \mathbf{d}_k{}^t(\mathbf{d}_k{}^{t'}),$$

or, since $\mathbf{d}_k{}^t$ is a vector whose elements $d_{nk}{}^t$ $(n=1, 2)$ equal one or zero,

$$PS^t(\mathbf{r},\mathbf{d}) = \mathbf{r}^t(\mathbf{r}^t)' - 2\mathbf{r}^t\big[(1/K^t) \sum_{k=1}^{K^t} (\mathbf{d}_k{}^t)'\big]$$

$$+ \big[(1/K^t) \sum_{k=1}^{K^t} \mathbf{d}_k{}^t\big]\mathbf{u}', \quad (6)$$

where $\mathbf{u}$ is a row vector whose elements both equal one, i.e., $\mathbf{u} = (1,1)$. Let

$$\mathbf{d}_k{}^t = \bar{\mathbf{d}}^t + (\mathbf{d}_k{}^t)^0,$$

where

$$\bar{\mathbf{d}}^t = (1/K^t) \sum_{k=1}^{K^t} \mathbf{d}_k{}^t,$$

and $(\mathbf{d}_k{}^t)^0$ is the difference between $\mathbf{d}_k{}^t$ and $\bar{\mathbf{d}}^t$. If we substitute $\bar{\mathbf{d}}^t$ into (6) and complete the square, then $PS^t(\mathbf{r},\mathbf{d})$ becomes

$$PS^t(\mathbf{r},\mathbf{d}) = (\mathbf{r}^t - \bar{\mathbf{d}}^t)(\mathbf{r}^t - \bar{\mathbf{d}}^t)' + \bar{\mathbf{d}}^t(\mathbf{u} - \bar{\mathbf{d}}^t)'. \quad (7)$$

Therefore, $PS(\mathbf{r},\mathbf{d})$, the weighted sum of the sub-collection scores, can be expressed as

$$PS(\mathbf{r},\mathbf{d}) = (1/K) \sum_{t=1}^{T} K^t(\mathbf{r}^t - \bar{\mathbf{d}}^t)(\mathbf{r}^t - \bar{\mathbf{d}}^t)'$$

$$+ (1/K) \sum_{t=1}^{T} K^t\bar{\mathbf{d}}^t(\mathbf{u} - \bar{\mathbf{d}}^t)'. \quad (8)$$

In Eq. (8), $PS(\mathbf{r},\mathbf{d})$ represents the vector partition of the PS. Note that the vector elements in the terms in (8) are analogous to the scalar elements in the terms in (4). Thus, we can refer to the two terms on the RHS of (8) as measures of reliability and resolution, respectively, when forecasts are considered to be vectors (see footnote 11). However, in Section 5 we show that the scalar and vector measures of these attributes are not equivalent, i.e., linearly related.

5. Scalar and vector partitions: A comparison

Hereafter, for comparative purposes, we denote the vector expression for the PS by $PS^*(\mathbf{r},\mathbf{d})$, where $PS^*(\mathbf{r},\mathbf{d}) = \frac{1}{2}PS(\mathbf{r},\mathbf{d})$. Note that the range of $PS^*(\mathbf{r},\mathbf{d})$ is the closed unit interval $[0,1]$.

The scalar partition of the PS, $PS(r,d)$ in (4), can be expressed as

$$PS(r,d) = (1/M)\big[\sum_{s=1}^{S} M^s(r^s)^2$$

$$- 2 \sum_{s=1}^{S} M^s r^s \bar{d}^s + \sum_{s=1}^{S} M^s(\bar{d}^s)^2\big]$$

$$+ (1/M)\big[\sum_{s=1}^{S} M^s \bar{d}^s - \sum_{s=1}^{S} M^s(\bar{d}^s)^2\big]. \quad (9)$$

We denote the two sets of terms on the RHS of (9) by S1 and S2, respectively, and the terms which constitute these sets by S11, S12, and S13 and S21 and S22, respectively. Thus,

$$PS(r,d) = S1 + S2, \quad (10)$$

where

$$S1 = S11 + S12 + S13, \quad (11)$$

and

$$S2 = S21 + S22. \quad (12)$$

Note that $S13 = -S22$. Thus, $PS(r,d)$, in (10), can also be expressed as

$$PS(r,d) = S11 + S12 + S21. \quad (13)$$

The vector partition of the PS, $PS^*(\mathbf{r},\mathbf{d})$, can be expressed as

$$PS^*(\mathbf{r},\mathbf{d}) = (1/2K) \sum_{t=1}^{T} K^t \sum_{n=1}^{2} (r_n{}^t - \bar{d}_n{}^t)^2$$

$$+ (1/2K) \sum_{t=1}^{T} K^t \sum_{n=1}^{2} \bar{d}_n{}^t(1 - \bar{d}_n{}^t) \quad (14a)$$

[see (5a) or (8)], or as

$$PS^*(\mathbf{r},\mathbf{d}) = (1/2K)\big[\sum_{t=1}^{T} K^t \sum_{n=1}^{2} (r_n{}^t)^2$$

$$- 2 \sum_{t=1}^{T} K^t \sum_{n=1}^{2} r_n{}^t\bar{d}_n{}^t + \sum_{t=1}^{T} K^t \sum_{n=1}^{2} (\bar{d}_n{}^t)^2\big]$$

$$+ (1/2K)\big[\sum_{t=1}^{T} K^t \sum_{n=1}^{2} \bar{d}_n{}^t - \sum_{t=1}^{T} K^t \sum_{n=1}^{2} (\bar{d}_n{}^t)^2\big]. \quad (14b)$$

We denote the two sets of terms on the RHS of (14b) by V1 and V2, respectively, and the terms which constitute these sets by V11, V12, and V13 and V21 and V22, respectively. Thus,

$$PS^*(\mathbf{r},\mathbf{d}) = V1 + V2, \quad (15)$$

where

$$V1 = V11 + V12 + V13, \quad (16)$$

and

$$V2 = V21 + V22. \quad (17)$$

Note that $V13 = -V22$. Thus, $PS^*(\mathbf{r},\mathbf{d})$, in (15), can also be expressed as

$$PS^*(\mathbf{r},\mathbf{d}) = V11 + V12 + V21. \quad (18)$$

We compare the expressions for the scalar partition $PS(r,d)$, in (9), and the vector partition $PS^*(\mathbf{r},\mathbf{d})$, in (14b), term by term in the Appendix. We show that S11 = V11, S12 = V12, and S21 = V21. Thus, from (13) and (18),

$$PS(r,d) = PS^*(\mathbf{r},\mathbf{d}). \quad (19)$$

We also show that the difference between S13 and V13,

i.e., V13—S13, is

$$V13-S13=(1/2K)\sum_{t,t^*}[K^tK^{t^*}/(K^t+K^{t^*})]$$

$$\times[(\bar{d}_1{}^t-\bar{d}_2{}^{t^*})^2+(\bar{d}_1{}^{t^*}-\bar{d}_2{}^t)^2], \quad (20)$$

where the summation on the RHS of (20) is taken over all pairs of subcollections t and t^* for which $\mathbf{r}^t+\mathbf{r}^{t^*}=\mathbf{u}$ $(t, t^*=1, \ldots, T)$ (see Appendix).

Note, from (20), that

$$V13-S13\geqslant 0. \quad (21)$$

Further, note that equality holds, in (21), only if either $K^t=0$ or $K^{t^*}=0$ or if $\bar{d}_1{}^t=\bar{d}_2{}^{t^*}$ (in which case, $\bar{d}_1{}^{t^*}=\bar{d}_2{}^t$) for all pairs of subcollections t and $t^*(\mathbf{r}^t+\mathbf{r}^{t^*}$ $=\mathbf{u}; t, t^*=1, \ldots, T)$. Finally, note, from (11) and (16), that

$$V1-S1\geqslant 0, \quad (22)$$

and from Eqs. (12) and (17), since S13 = —S22 and

TABLE 1. A sample collection of forecasts and the relevant observations when forecasts and observations are considered to be (a) scalars and (b) vectors.

a. Scalar forecasts and observations

Forecast or observation number m	Forecast r_m	Observation d_m
1	0.2	0
2	0.8	1
3	0.6	1
4	0.4	0
5	0.9	1
6	0.1	0
7	0.2	0
8	0.8	1
9	0.1	0
10	0.9	1
11	0.2	0
12	0.8	1
13	0.4	1
14	0.6	0
15	0.7	1
16	0.3	0
17	0.8	1
18	0.2	0
19	0.2	1
20	0.8	0

b. Vector forecasts and observations

Forecast or observation number k	Forecast $\mathbf{r}_k$	Observation $\mathbf{d}_k$
1	(0.2, 0.8)	(0, 1)
2	(0.6, 0.4)	(1, 0)
3	(0.9, 0.1)	(1, 0)
4	(0.2, 0.8)	(0, 1)
5	(0.1, 0.9)	(0, 1)
6	(0.2, 0.8)	(0, 1)
7	(0.4, 0.6)	(1, 0)
8	(0.7, 0.3)	(1, 0)
9	(0.8, 0.2)	(1, 0)
10	(0.2, 0.8)	(1, 0)

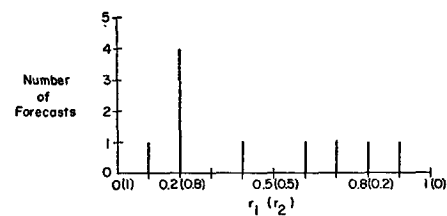(a) Scalar forecasts
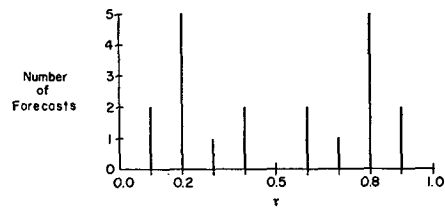
(b) Vector forecasts

FIG. 2. The sample collection of forecasts presented in Table 1 depicted within the appropriate framework when forecasts are considered to be (a) scalars and (b) vectors.

$V13=-V22$, that

$$V2-S2\leqslant 0. \quad (23)$$

Thus, the value of the reliability term for scalar forecasts is, in general, less than that for vector forecasts, while the value of the resolution term for scalar forecasts is, in general, greater than that for vector forecasts. That is, if a collection of forecasts is considered to consist of scalar forecasts, then the collection will appear, in general, to have more reliability and less resolution than if the collection is considered to consist of vector forecasts.

6. Scalar and vector partitions: A sample collection of forecasts

In order to illustrate the differences between the scalar and vector partitions of the PS, we consider a sample collection of probability forecasts. The forecasts and the relevant observations are presented in Tables 1a and 1b, in which we identify these quantities as scalars and vectors, respectively. We depict the forecasts as scalars and vectors within the appropriate framework in Figs. 2a and 2b, respectively.

The scalar partition for these forecasts is presented in Table 2. Note that the values of the terms S1 (reliability) and S2 (resolution) are 0.013 and 0.130, respectively, and that their sum, i.e., $PS(r,d)$, equals 0.143. The vector partition for these forecasts is presented in Table 3. Note that the values of the terms V1 (reliability) and V2 (resolution) are 0.068 and 0.075, respectively, and that their sum, i.e., $PS^*(\mathbf{r},\mathbf{d})$, also

TABLE 2. The scalar partition for the sample collection of forecasts presented in Table 1a.

Subcollection number s	Forecast r^s	Number of forecasts M^s	Observed relative frequency d^s	Kernel reliability $(r^s-d^s)^2$	Kernel resolution $d^s(1-d^s)$	Subcollection reliability S1(s) $M^s(r^s-d^s)^2$	Subcollection resolution S2(s) $M^s d^s(1-d^s)$
1	0.1	2	0.0	0.01	0.00	0.02	0.00
2	0.2	5	0.2	0.00	0.16	0.00	0.80
3	0.3	1	0.0	0.09	0.00	0.09	0.00
4	0.4	2	0.5	0.01	0.25	0.02	0.50
5	0.6	2	0.5	0.01	0.25	0.02	0.50
6	0.7	1	1.0	0.09	0.00	0.09	0.00
7	0.8	5	0.8	0.00	0.16	0.00	0.80
8	0.9	2	1.0	0.01	0.00	0.02	0.00
Total		20				0.26	2.60
Average						0.013	0.130

TABLE 3. The vector partition for the sample collection of forecasts presented in Table 1b.

Sub-collection number t	Forecast $\mathbf{r}^t$	Number of forecasts K^t	Observed relative frequency $\bar{\mathbf{d}}^t$	Kernel reliability $(\mathbf{r}^t-\bar{\mathbf{d}}^t)(\mathbf{r}^t-\bar{\mathbf{d}}^t)'$	Kernel resolution $\bar{\mathbf{d}}^t(\mathbf{u}-\bar{\mathbf{d}}^t)'$	Subcollection reliability V1(t) $K^t(\mathbf{r}^t-\bar{\mathbf{d}}^t)(\mathbf{r}^t-\bar{\mathbf{d}}^t)'$	Subcollection resolution V2(t) $K^t\bar{\mathbf{d}}^t(\mathbf{u}-\bar{\mathbf{d}}^t)'$
1	(0.1, 0.9)	1	(0.00, 1.00)	0.020	0.000	0.02	0.00
2	(0.2, 0.8)	4	(0.25, 0.75)	0.005	0.375	0.02	1.50
3	(0.4, 0.6)	1	(1.00, 0.00)	0.720	0.000	0.72	0.00
4	(0.6, 0.4)	1	(1.00, 0.00)	0.320	0.000	0.32	0.00
5	(0.7, 0.3)	1	(1.00, 0.00)	0.180	0.000	0.18	0.00
6	(0.8, 0.2)	1	(1.00, 0.00)	0.080	0.000	0.08	0.00
7	(0.9, 0.1)	1	(1.00, 0.00)	0.020	0.000	0.02	0.00
Total		10				1.36	1.50
Average*						0.068	0.075

* This average is computed on the basis of $2K = 20$ forecasts [see (14a)].

equals 0.143. Thus, as indicated in (22) and (23), V1(0.068)$\geqslant$S1(0.013) and V2(0.075)$\leqslant$S2(0.130). Note that for this collection of forecasts the resolution term is ten times as large as the reliability term according to the scalar partition, while the resolution term is only slightly larger than the reliability term according to the vector partition. Thus, while the differences between the terms in these partitions can be expected, in general, to decrease as the number of forecasts in a collection increases, substantial differences may occur, at least for small collections.

7. The use of scalar and vector partitions: Discussion

A number of questions arise in connection with the use of scalar and vector partitions of the PS. For example: 1) Which partitions have been used by evaluators in previous forecast evaluation studies? 2) How sensitive are the results of these studies to the particular partition used? 3) Which partition should an evaluator use in such studies? 4) What are the effects of the definitions of scalar and vector forecasts upon the partitions and their use?

With regard to the first question, evaluators, in general, have used a special "scalar" partition, a partition which is based upon an expression for the PS

in which only the probabilities assigned to one of the two states are considered.[12] If we denote this expression for the PS by $PS^*(r,d)$, then, from (5a),

$$PS^*(r,d) = (1/K) \sum_{k=1}^{K} (r_{nk}-d_{nk})^2, \quad n=1,2. \tag{24}$$

Thus, the special scalar partition can be expressed as

$$PS^*(r,d) = (1/K) \sum_{t=1}^{T} K^t(r_n{}^t-\bar{d}_n{}^t)^2$$

$$+(1/K) \sum_{t=1}^{T} K^t\bar{d}_n{}^t(1-\bar{d}_n{}^t),$$

or, since

$$\sum_{n} (r_n{}^t-\bar{d}_n{}^t)^2 = 2(r_n{}^t-\bar{d}_n{}^t)^2$$

[12] A determination of which partition has been used in a particular study cannot always be made. However, for those studies for which such a determination could be made, the special scalar partition, in general, has been the partition used. Why have evaluators (including Sanders) used this partition instead of the scalar partition proposed by Sanders (see footnote 3)? We can only assume that they believed that these partitions were equivalent and, as a result, used the special scalar partition because this partition appeared to be easier to apply.

TABLE 4. The special scalar partition for the sample collection of forecasts presented in Table 1 in terms of the probabilities assigned to state s_1.

Subcollection number t	Forecast $r_1{}^t$	Number of forecasts $K_1{}^t$	Observed relative frequency $\bar{d}_1{}^t$	Kernel reliability $(r_1{}^t - \bar{d}_1{}^t)^2$	Kernel resolution $\bar{d}_1{}^t(1 - \bar{d}_1{}^t)$	Subcollection reliability $S1^*(t)$ $K^t(r_1{}^t - \bar{d}_1{}^t)^2$	Subcollection resolution $S2^*(t)$ $K^t\bar{d}_1{}^t(1 - \bar{d}_1{}^t)$
1	0.1	1	0.00	0.0100	0.0000	0.01	0.00
2	0.2	4	0.25	0.0025	0.1875	0.01	0.75
3	0.4	1	1.00	0.3600	0.0000	0.36	0.00
4	0.6	1	1.00	0.1600	0.0000	0.16	0.00
5	0.7	1	1.00	0.0900	0.0000	0.09	0.00
6	0.8	1	1.00	0.0400	0.0000	0.04	0.00
7	0.9	1	1.00	0.0100	0.0000	0.01	0.00
Total		10				0.68	0.75
Average						0.068	0.075

and

$$\sum_n \bar{d}_n{}^t(1 - \bar{d}_n{}^t) = 2\bar{d}_n{}^t(1 - \bar{d}_n{}^t), \quad n = 1, 2,$$

$$PS^*(r,d) = (1/2K) \sum_{t=1}^{T} K^t \sum_{n=1}^{2} (r_n{}^t - \bar{d}_n{}^t)^2$$

$$+ (1/2K) \sum_{t=1}^{T} K^t \sum_{n=1}^{2} \bar{d}_n{}^t(1 - \bar{d}_n{}^t). \quad (25)$$

Note that the terms on the RHS of (25) are identical to the terms on the RHS of (14a). Thus, the special scalar partition is equivalent to the vector partition. We present the special scalar partition for the sample collection of forecasts (see Table 1) in Table 4 in terms of the probabilities assigned to state s_1.

We are aware of only one study, that by Sanders (1958, 1963), in which another partition, the scalar partition, has been used. In his study, which involved forecasts for many different variables, some of which were two-state $(N = 2)$ variables and some of which were N-state $(N > 2)$ variables, Sanders considered only one of the two states for the two-state variables and each state separately for the N-state variables. Thus, Sanders applied a special scalar, or vector, partition to the two-state forecasts and a scalar partition to the N-state forecasts.

With regard to the second question, we have not, as yet, applied the scalar and vector partitions to any large collections of forecasts in order to determine the differences between the respective measures of reliability and resolution. However, as indicated in Section 6, these differences can be substantial for small collections of forecasts. In this regard, although we are not able to determine the magnitude of this difference in Sanders' study, the reliability (resolution) of the forecasts in his study would certainly have decreased (increased) if the vector partition had been applied to the N-state $(N > 2)$ forecasts as well as to the two-state $(N = 2)$ forecasts (see Murphy, 1971b). Finally, since we believe that the vector partition is, in general, the more appropriate partition (see below), we should indicate that, since the studies conducted heretofore have been concerned primarily with two-state variables and since the evaluators in most, if not all, of these studies have used the special scalar partition, the results obtained in these studies are equivalent to the results that would have been obtained if the vector partition had been used.

The question of which partition an evaluator should use can be considered from at least two points of view. From a "scientific" point of view the answer to this question depends, in part, upon whether the evaluator, a meteorologist, is concerned with the reliability and resolution of *forecasts* or *probabilities*. If forecasts are of primary concern, then, since forecasts are vectors, the vector partition must be used to determine their reliability and resolution, while, if the probabilities which constitute forecasts are of primary concern, then, since probabilities are scalars, the scalar partition must be used. In addition, the vector partition is, and the scalar partition is not, formulated in such a way that a one-to-one correspondence exists between the probabilities and the states for each subcollection of forecasts.[13] Since we believe that meteorologists are, or should be, primarily concerned with the reliability and resolution of forecasts and that, whenever possible, a one-to-one correspondence should be maintained between the probabilities and the states, the vector partition appears to be more appropriate than the scalar partition from a scientific point of view.

From an "economic" point of view the evaluator is a decision maker, i.e., a user of probability forecasts, in a two-state $(N = 2)$ decision situation. The decision maker will be particularly concerned with the reliability and resolution of forecasts in the vicinity of his "indifference" points, i.e., the points, or forecasts, for which he is indifferent between two actions (see Murphy, 1971a). Since the unit line segment with a length coordinate system (see Section 2) represents the proper framework within which to depict these indifference points, only the vector partition can provide the decision

---

[13] This reason for preferring the vector partition to the scalar partition was brought to the author's attention by Chien-hsiung Yang.

maker with the appropriate information. For example, consider a decision maker in a two-action, two-state decision situation who is indifferent between these actions when the probability of precipitation is 0.2 (and, as a result, the probability of no precipitation is 0.8). This decision maker, then, will be particularly concerned with the reliability and resolution of forecasts in the vicinity of the point (0.2,0.8). On the other hand, he will not be particularly concerned with either the reliability and resolution of forecasts in the vicinity of the point (0.8, 0.2) or the composite reliability and resolution of the probabilities 0.2 and 0.8 (the scalar partition would provide the latter). Thus, the vector partition appears to be more appropriate than the scalar partition from an economic as well as a scientific point of view.

With regard to the fourth question, we consider only the effect of the definitions of scalar and vector forecasts upon "sample size" in this paper. In the two-state ($N=2$) situation the number of distinct scalar forecasts and the number of distinct vector forecasts are both equal to S, the number of distinct probability values (see footnote 9). However, since a collection of forecasts consists of M scalar and K ($=M/2$) vector forecasts, the sample size in the scalar framework is twice that in the vector framework. Thus, an evaluator who is concerned with obtaining estimates of the reliability and resolution of certain distinct forecasts will have, on the average, twice as many scalar as vector forecasts upon which to base these estimates. Therefore, for small collections of forecasts the number of vector forecasts may not be sufficient to obtain reasonable estimates of these attributes for certain forecasts. One possible solution to this problem would be to combine those subcollections which correspond to "adjacent" forecasts with the subcollection which corresponds to the forecast of concern. Such a procedure can be expected to provide reasonable estimates of these attributes for most, if not all, vector forecasts. We discuss this problem in greater detail in Murphy (1971b).[14]

8. Conclusion

In this paper we have described and compared scalar and vector partitions of the PS in the two-state ($N=2$) situation. These partitions, which are based upon expressions for the PS in which probability forecasts are considered to be scalars and vectors, respectively, provide similar, but not equivalent (i.e., linearly related), measures of the reliability and resolution of the forecasts. Specifically, the reliability (resolution) of the forecasts according to the scalar partition is, in general, greater (less) than their reliability (resolution) according to the vector partition. A sample collection

of forecasts has been used to illustrate the differences between these partitions.

We have briefly considered several questions related to the use of these partitions in the two-state ($N=2$) situation. In this regard, we have indicated that: 1) a special scalar partition, which is equivalent to the vector partition in the two-state situation, has been used in most, if not all, forecast evaluation studies; 2) the differences between the reliability and resolution terms in these partitions can be substantial, at least for small collections of forecasts; 3) the vector partition appears to be more appropriate than the scalar partition from a scientific as well as an economic point of view; and 4) the use of the vector, rather than the scalar, partition "reduces" the effective sample size of a collection of forecasts.

As indicated in Section 2, the differences between the scalar and vector partitions of the PS become more evident in N-state ($N>2$) situations. We describe and compare these partitions in N-state situations in a separate paper (see Murphy, 1971b).

APPENDIX

We compare the scalar and vector partitions of the PS, $PS(r,d)$, in (9), and $PS^*(\mathbf{r},\mathbf{d})$, in (14b), respectively, term by term in this Appendix. Since S13 $=-$S22 and V13 $=-$V22, we compare only S11 and V11, S12 and V12, S13 and V13, and S21 and V21. Further, since $M=2K$, the coefficients of the respective terms are equal and, as a result, they need not be considered.

We consider only the elements in the scalar terms which correspond to the subcollections s and s^*, for which $r_m=r^s$ and $r_m=r^{s*}$, respectively, where

$$r^{s*}=1-r^s, \qquad (\text{A1})$$

and we consider only the elements in the vector terms which correspond to the subcollections t and t^*, for which $\mathbf{r}_k=\mathbf{r}^t$ and $\mathbf{r}_k=\mathbf{r}^{t*}$, respectively, where

$$\mathbf{r}^{t*}=\mathbf{u}-\mathbf{r}^t. \qquad (\text{A2})$$

The comparison of the sums of these elements is sufficient to determine the relationship between the respective terms, because a collection of scalar or vector forecasts consists of such pairs of subcollections of forecasts.

Since the number S of distinct scalar forecasts and the number T of distinct vector forecasts are equal in the two-state ($N=2$) situation (see Section 2), we can assume, without any loss of generality, that the subcollection s of scalar forecasts for which $r_m=r^s$ and the subcollection t of vector forecasts for which $\mathbf{r}_k=\mathbf{r}^t$ correspond in such a way that

$$r^s=r_1{}^t \quad (s=1, \ldots, S; t=1, \ldots, T). \qquad (\text{A3})$$

Then, since $r_1{}^t+r_2{}^t=1$,

$$r_2{}^t=1-r^s, \qquad (\text{A4})$$

---

[14] This problem is of greater concern in N-state ($N>2$) situations, in which the number T of distinct vector forecasts is greater than the number S of distinct scalar forecasts (see footnote 9).

As a result of the correspondence between the sub-collections of scalar and vector forecasts, note that

$$M^s = K^t + K^{t*}, \tag{A5}$$

$$M^{s*} = K^t + K^{t*}, \tag{A6}$$

and that

$$\bar{d}^s = [1/(K^t+K^{t*})](K^t\bar{d}_1{}^t + K^{t*}\bar{d}_2{}^{t*}), \tag{A7}$$

$$\bar{d}^{s*} = [1/(K^t+K^{t*})](K^{t*}\bar{d}_1{}^{t*} + K^t\bar{d}_2{}^t). \tag{A8}$$

a. S11 and V11

Let $S11(s,s^*)$ denote the sum of the two elements of concern in the scalar term. Then,

$$S11(s,s^*) = M^s(r^s)^2 + M^{s*}(r^{s*})^2,$$

or, from (A1) and (A3)–(A6),

$$S11(s,s^*) = (K^t + K^{t*})[(r_1{}^t)^2 + (r_2{}^t)^2].$$

Let $V11(t,t^*)$ denote the sum of the two elements of concern in the vector term. Then

$$V11(t,t^*) = K^t \sum_{n=1}^{2} (r_n{}^t)^2 + K^{t*} \sum_{n=1}^{2} (r_n{}^{t*})^2,$$

or

$$V11(t,t^*) = K^t[(r_1{}^t)^2 + (r_2{}^t)^2] + K^{t*}[(r_1{}^{t*})^2 + (r_2{}^{t*})^2],$$

or, from (A2),

$$V11(t,t^*) = (K^t + K^{t*})[(r_1{}^t)^2 + (r_2{}^t)^2].$$

Thus,

$$S11(s,s^*) = V11(t,t^*),$$

and, as a result,

$$S11 = V11.$$

b. S12 and V12

Let $S12(s,s^*)$ denote the sum of the two elements of concern in the scalar term. Then,

$$S12(s,s^*) = M^s r^s \bar{d}^s + M^{s*} r^{s*} \bar{d}^{s*},$$

or, from (A1)–(A8),

$$S12(s,s^*) = K^t r_1{}^t \bar{d}_1{}^t + K^{t*} r_2{}^{t*} \bar{d}_2{}^{t*} + K^{t*} r_1{}^{t*} \bar{d}_1{}^{t*} + K^t r_2{}^t \bar{d}_2{}^t.$$

Let $V12(t,t^*)$ denote the sum of the two elements of concern in the vector term. Then,

$$V12(t,t^*) = K^t \sum_{n=1}^{2} r_n{}^t \bar{d}_n{}^t + K^{t*} \sum_{n=1}^{2} r_n{}^{t*} \bar{d}_n{}^{t*},$$

or

$$V12(t,t^*) = K^t r_1{}^t \bar{d}_1{}^t + K^t r_2{}^t \bar{d}_2{}^t + K^{t*} r_1{}^{t*} \bar{d}_1{}^{t*} + K^{t*} r_2{}^{t*} \bar{d}_2{}^{t*}.$$

Thus,

$$S12(s,s^*) = V12(t,t^*),$$

and, as a result,

$$S12 = V12.$$

c. S13 and V13

Let $S13(s,s^*)$ denote the sum of the two elements of concern in the scalar term. Then,

$$S13(s,s^*) = M^s(\bar{d}^s)^2 + M^{s*}(\bar{d}^{s*})^2,$$

or, from (A5)–(A8),

$$S13(s,s^*) = [1/(K^t+K^{t*})][(K^t\bar{d}_1{}^t)^2 + 2K^tK^{t*}\bar{d}_1{}^t\bar{d}_2{}^{t*} + (K^{t*}\bar{d}_2{}^{t*})^2] + [1/(K^t+K^{t*})][(K^{t*}\bar{d}_1{}^{t*})^2 + 2K^tK^{t*}\bar{d}_1{}^{t*}\bar{d}_2{}^t + (K^t\bar{d}_2{}^t)^2].$$

Let $V13(t,t^*)$ denote the sum of the two elements of concern in the vector term. Then,

$$V13(t,t^*) = K^t \sum_{n=1}^{2} (\bar{d}_n{}^t)^2 + K^{t*} \sum_{n=1}^{2} (\bar{d}_n{}^{t*})^2,$$

or

$$V13(t,t^*) = K^t[(\bar{d}_1{}^t)^2 + (\bar{d}_2{}^t)^2] + K^{t*}[(\bar{d}_1{}^{t*})^2 + (\bar{d}_2{}^{t*})^2],$$

or

$$V13(t,t^*) = [1/(K^t+K^{t*})][(K^t\bar{d}_1{}^t)^2 + K^tK^{t*}(\bar{d}_1{}^t)^2 + (K^t\bar{d}_2{}^t)^2 + K^tK^{t*}(\bar{d}_2{}^t)^2] + [1/(K^t+K^{t*})][(K^{t*}\bar{d}_1{}^{t*})^2 + K^tK^{t*}(\bar{d}_1{}^{t*})^2 + (K^{t*}\bar{d}_2{}^{t*})^2 + K^tK^{t*}(\bar{d}_2{}^{t*})^2].$$

The difference between $S13(s,s^*)$ and $V13(t,t^*)$ is then $V13(t,t^*) - S13(s,s^*)$, where

$$V13(t,t^*) - S13(s,s^*) = [1/(K^t+K^{t*})][K^tK^{t*}(\bar{d}_1{}^t)^2 - 2K^tK^{t*}\bar{d}_1{}^t\bar{d}_2{}^{t*} + K^tK^{t*}(\bar{d}_2{}^{t*})^2] + [1/(K^t+K^{t*})] \times [K^tK^{t*}(\bar{d}_1{}^{t*})^2 - 2K^tK^{t*}\bar{d}_1{}^{t*}\bar{d}_2{}^t + K^tK^{t*}(\bar{d}_2{}^t)^2],$$

or

$$V13(t,t^*) - S13(s,s^*) = [K^tK^{t*}/(K^t+K^{t*})] \times [(\bar{d}_1{}^t - \bar{d}_2{}^{t*})^2 + (\bar{d}_1{}^{t*} - \bar{d}_2{}^t)^2].$$

Thus,

$$V13 - S13 = (1/2K) \sum_{t,t^*} [K^tK^{t*}/(K^t+K^{t*})] \times [(\bar{d}_1{}^t - \bar{d}_2{}^{t*})^2 + (\bar{d}_1{}^{t*} - \bar{d}_2{}^t)^2],$$

where the summation is taken over all pairs of sub-collections t and t^* for which $\mathbf{r}^t + \mathbf{r}^{t*} = \mathbf{u}(t, t^* = 1, \ldots, T)$.

d. S21 and V21

Let $S21(s,s^*)$ denote the sum of the two elements of concern in the scalar term. Then,

$$S21(s,s^*) = M^s \bar{d}^s + M^{s*} \bar{d}^{s*},$$

or, from (A5)–(A8),

$$S21(s,s^*) = K^t\bar{d}_1{}^t + K^{t*}\bar{d}_2{}^{t*} + K^{t*}\bar{d}_1{}^{t*} + K^t\bar{d}_2{}^t.$$

Let $V21(t,t^*)$ denote the sum of the two elements of concern in the vector term. Then,

$$V21(t,t^*) = K^t \sum_{n=1}^{2} \bar{d}_n{}^t + K^{t*} \sum_{n=1}^{2} \bar{d}_n{}^{t*},$$

or

$$V21(t,t^*) = K^t \bar{d}_1{}^t + K^t \bar{d}_2{}^t + K^{t*} \bar{d}_1{}^{t*} + K^{t*} \bar{d}_2{}^{t*}.$$

Thus,

$$S21(s,s^*) = V21(t,t^*),$$

and, as a result,

$$S21 = V21.$$

REFERENCES

Brier, G. W., 1950: Verification of forecasts expressed in terms of probability. *Mon. Wea. Rev.*, **78**, 1–3.

Bross, I. D. J., 1953: *Design for Decision.* New York, MacMillian, 276 pp.

Murphy, A. H., 1971a: Ordinal relationships between measures of the "accuracy" and "value" of probability forecasts: Preliminary results. Unpubl. ms., Dept. of Meteorology and Oceanography, University of Michigan, 52 pp.

——, 1971b: Scalar and vector partitions of the probability score: Part II, *N*-state situation. Unpubl. ms., National Center for Atmospheric Research, Boulder, Colo., 47 pp.

——, and E. S. Epstein, 1967: Verification of probabilistic predictions: A brief review. *J. Appl. Meteor.*, **6**, 748–755.

——, and R. L. Winkler, 1970: Scoring rules in probability assessment and evaluation. *Acta Psychologica*, **34**, 273–286.

Pontryagin, L. S., 1952: *Foundations of Combinatorial Topology.* Rochester, N. Y., Graylock Press, 99 pp.

Sanders, F., 1958: The evaluation of subjective probability forecasts. Tech. Rept. No. 5, Contract AF 19(604)-1305, Dept. of Meteorology, MIT, 62 pp.

——, 1963: On subjective probability forecasting. *J. Appl. Meteor.*, **2**, 191–201.

Springer, C. E., 1964: *Geometry and Analysis of Projective Spaces.* San Francisco, W. H. Freeman, 299 pp.

Winkler, R. L., and A. H. Murphy, 1968: "Good" probability assessors. *J. Appl. Meteor.*, **7**, 751–758.